

# 以輪廓為特徵之書法文字書寫者辨識

詹明華<sup>1\*</sup>、鄒濬智<sup>2</sup>、何冠緯<sup>3</sup>

<sup>1</sup> 中央警察大學刑事警察學系

<sup>2</sup> 中央警察大學通識中心

<sup>3</sup> 桃園市政府警察局刑事鑑識中心

\*E-mail: mhchan@mail.cpu.edu.tw

## 摘要

傳統上，書法文字的書寫者辨識，係以專家的知識為基礎來進行判斷，惟此種判斷方式較為主觀，倘若結果不被他人所接受，往往會演變成各持己見爭執不下的窘境，若能建立電腦化的科學辨識，則辨識結果較為客觀，而且電腦化的科學辨識會不斷地進步，辨識的正確率會不斷地提升。本文以書法文字的輪廓為特徵，建立了一套電腦化的辨識系統，首先擷取文字的輪廓，然後計算輪廓的斜率，由斜率間的距離，可分辨出該文字的書寫者。實驗的對象為9個字，書寫者共有3人，每人提供3個字，由初步的實驗結果顯示，系統可將9個字分成3群，且每一群均有3個字，即系統可正確地辨識出此9個字為3人書寫，且每人書寫3個字。

**關鍵字：**書法文字、輪廓、電腦化辨識

## 前言

傳統上，書法文字的書寫者辨識，係以專家的知識與經驗為基礎來進行判斷。一般專家會運用顯微鏡與測量器，以人工的方式來計算與分析[1]，此種人工的方式非常耗時費力[2]，加上判斷方式較為主觀，倘若結果不被他人所接受，往往會演變成各持己見爭執不下的窘境。專家的知識與經驗會隨著人的死亡而消失，傳承不易也無法複製，但電腦卻能不斷地累積經驗且複製容易。因此，利用電腦科技來發展書法文字的辨識，為一值得努力的方向。

## 辨識特徵與演算法

書法文字係由各種筆劃構成，我們可用各種斜率的直線來近似各種筆劃的輪廓。若有兩點 $(x_1, y_1)$ 與 $(x_2, y_2)$ 位於一條直線上，則此直線的斜率為

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad (1)$$

假設有 $p$ 個書法文字欲辨識其書寫者，我們以某一筆劃的輪廓為特徵進行辨識，若此一筆劃可用 $n$ 條直線近似其輪廓，則我們可用 $pn$ 個斜率為特徵來辨識書寫者。假設第 $i$ 個文字第 $j$ 條直線的斜率為 $m_{ij}$ ,  $i=1, 2, \dots, p$ ,  $j=1, 2, \dots, n$ ，在 $p$ 個文字中，若第 $j$ 個斜率的最小值為 $a_j$ ，則第 $j$ 個斜率可正規化為

$$m_{ij}' = m_{ij} - a_j \quad (2)$$

經過(2)式的計算後，每一種斜率的最小值均為0。經過(2)式處理後，若第 $j$ 個斜率的最大值為 $b_j$ ，則第 $j$ 個斜率可正規化為

$$m_{ij}'' = m_{ij}' / b_j \quad (3)$$

經過(3)式的計算後，每一種斜率的最小值均為0且最大值均為1。在 $p$ 個文字中，第 $k$ 個字與第 $l$ 個字的斜率總距離可定義為

$$d_{kl} = \sum_{j=1}^n |m_{kj}'' - m_{lj}''| \quad (4)$$

利用(4)式，可算出每一個文字與其他文字的所有斜率總距離。接下來，透過斜率總距離，進行文字書寫者的辨識。我們將分群的斜率總距離門檻值設為 $t$ ，即在相同群中，文字間的斜率

總距離不能超過  $t$ ，經由不斷地組合測試，最後可得到分群的結果，即算出有多少書寫者與每一書寫者所擁有的文字。

### 結果與討論

如圖一所示 9 個書法古文字，由 3 個人提供，每人各提供 3 個字，每一列的 3 個字均為相同人所書寫。首先，將每一個字最後的筆劃抽離，每一個字對應的筆劃如圖二所示。對於圖二的筆劃，其主要的書寫特徵為落筆處與收筆處，本文只針對落筆處進行識別，如圖三所示，可用三條直線近似圖二每一筆劃的落筆處之輪廓，對於第一個筆劃，其下輪廓、右輪廓與上輪廓的斜率分別為 0.5217、-0.8723 與 0.0000，其餘八個筆劃的斜率如表一所示。

由表一可知，下輪廓、右輪廓與上輪廓的斜率最小值分別為 0.1579、-1.2381 與 0.0000，將下輪廓的斜率全部減 0.1579，將右輪廓的斜率全部加 -1.2381，將上輪廓的斜率全部減 0.0000，可將輪廓的最小斜率正規化為 0，如表二所示。由表二可知，下輪廓、右輪廓與上輪廓的斜率最大值分別為 0.7796、0.7619 與 0.3273，將下輪廓的斜率全部除以 0.7796，將右輪廓的斜率全部除以 0.7619，將上輪廓的斜率全部除以 0.3273，可將每一輪廓的斜率正規化為 0~1，如表三所示。

由表三可知，編號一的下輪廓、右輪廓與上輪廓之斜率分別為 0.4667、0.4801 與 0.0000，編號二的下輪廓、右輪廓與上輪廓之斜率分別為 1.0000、0.7182 與 0.2152，編號一與編號二的下輪廓、右輪廓與上輪廓之斜率距離分別為  $1.0000-0.4667=0.5333$ 、 $0.7182-0.4801=0.2381$  與  $0.2152-0.0000=0.2152$ ，編號一與編號二的斜率總距離為  $0.5333+0.2381+0.2152=0.9866$ ，以相同的方式，可求得編號一與其他編號的斜率總距離，如表四所示。接下來，再以相同的方式，可求得其他編號間的斜率總距離，如表四所示。

接下來，透過斜率總距離，進行文字書寫者的辨識。我們將分群的斜率總距離門檻值設為 1.000，即在相同群中，編號間的斜率總距離不能超過 1.000。表五為滿足條件的第一種分群結果，共有 ABC 三群，A 群中有編號一、二與三共 3 個成員，B 群中有編號四、五與六共 3 個成員，C 群中有編號七、八與九共 3 個成員，此為正確的分群結果。表六為滿足條件的第二種分群結果，共有 ABC 三群，A 群中有編號一、二與三共 3 個成員，B 群中只有編號四的 1 個成員，C 群中有編號五、六、七、八與九共 5 個成員，此為錯誤的分群結果，觀察圖三筆劃的落筆處可知，編號五、六、七、八與九雖為不同人所書寫，其輪廓卻極為近似，編號四、五與六雖為相同人所書寫，編號四卻略有不同。雖然表六為錯誤的分群結果，將來辨識時，我們可加入更多的特徵，應可將此錯誤排除。

### 結論

本文以書法文字的輪廓為特徵，建立了一套電腦化的辨識系統，首先擷取文字的輪廓，然後計算輪廓的斜率，由斜率間的距離，可分辨出該文字的書寫者，由初步的實驗結果顯示，系統可正確地辨識出書寫者。書法文字係人所書寫，書寫每一筆劃時很難一致，造成變異性，其書寫者的辨識充滿挑戰性，將來系統可加入專家的意見，找出更具識別度的特徵，以提高系統辨識的準確性，相信電腦化的科學辨識會不斷地進步，辨識的正確率會不斷地提升。

### 誌謝

本文書法古文字係由蔡勝宇、李豐碩與陳采容三位同學所提供，謹此誌謝。

### 參考文獻

- [1] 謝瑞智，民國 88 年，警察百科全書(十二)刑事鑑識，中央警察大學印行，桃園，頁 193-225。
- [2] 陳虎生，民國 85 年，筆跡鑑定之探討，警學叢刊，第 26 卷第 5 期，頁 17-29。



(一)

(二)

(三)



(四)



(五)



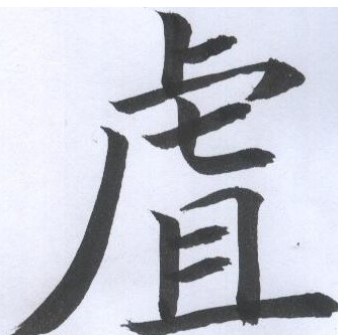
(六)



(七)



(八)



(九)

圖一、由 3 個人各提供 3 個書法古文字



(一)



(二)



(三)



(四)



(五)



(六)



(七)

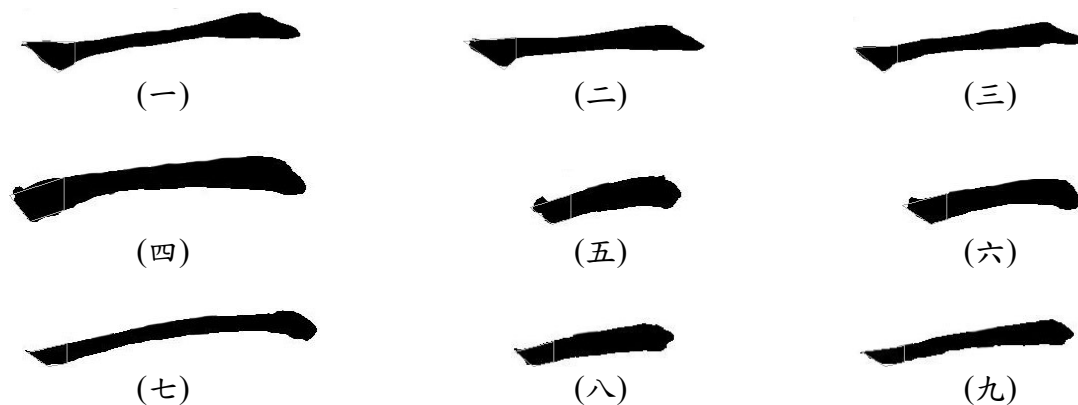


(八)



(九)

圖二、圖一每一文字對應的最後筆劃



圖三、以三條直線近似圖二每一筆劃的落筆處之輪廓

表一、圖三輪廓的斜率

編號	一	二	三	四	五	六	七	八	九
下輪廓	0.5217	0.9375	0.7778	0.3235	0.3158	0.3125	0.1579	0.2105	0.2105
右輪廓	-0.8723	-0.6909	-0.8000	-1.2381	-0.7727	-0.6552	-0.4762	-0.7778	-0.6087
上輪廓	0.0000	0.0704	0.0517	0.3273	0.3171	0.2667	0.2250	0.2703	0.1429

表二、將每一輪廓的最小斜率正規化為 0

編號	一	二	三	四	五	六	七	八	九
下輪廓	0.3638	0.7796	0.6199	0.1656	0.1579	0.1546	0.0000	0.0526	0.0526
右輪廓	0.3658	0.5472	0.4381	0.0000	0.4654	0.5829	0.7619	0.4603	0.6294
上輪廓	0.0000	0.0704	0.0517	0.3273	0.3171	0.2667	0.2250	0.2258	0.1429

表三、將每一輪廓的斜率正規化為 0~1

編號	一	二	三	四	五	六	七	八	九
下輪廓	0.4667	1.0000	0.7951	0.2125	0.2025	0.1983	0.0000	0.0675	0.0675
右輪廓	0.4801	0.7182	0.5750	0.0000	0.6108	0.7651	1.0000	0.6042	0.8261
上輪廓	0.0000	0.2152	0.1580	1.0000	0.9688	0.8148	0.6875	0.6900	0.4365

表四、每一筆劃間的斜率總距離

		一	二	三	四	五	六	七	八	九
一	下輪廓	0.0000	0.5333	0.3284	0.2542	0.2642	0.2684	0.4667	0.3992	0.3992
	右輪廓	0.0000	0.2381	0.0949	0.4801	0.1307	0.2850	0.5199	0.1241	0.3460
	上輪廓	0.0000	0.2152	0.1580	1.0000	0.9688	0.8148	0.6875	0.6900	0.4365
	總距離	0.0000	0.9866	0.5814	1.7343	1.3637	1.3682	1.6742	1.2133	1.1817
二	下輪廓	0.5333	0.0000	0.2049	0.7875	0.7975	0.8017	1.0000	0.9325	0.9325
	右輪廓	0.2381	0.0000	0.1432	0.7182	0.1074	0.0469	0.2818	0.1140	0.1079
	上輪廓	0.2152	0.0000	0.0571	0.7848	0.7537	0.5996	0.4723	0.4748	0.2213
	總距離	0.9866	0.0000	0.4052	2.2905	1.6585	1.4482	1.7541	1.5213	1.2617
三	下輪廓	0.3284	0.2049	0.0000	0.5827	0.5926	0.5968	0.7951	0.7276	0.7276
	右輪廓	0.0949	0.1432	0.0000	0.5750	0.0358	0.1901	0.4250	0.0292	0.2511
	上輪廓	0.1580	0.0571	0.0000	0.8420	0.8108	0.6568	0.5295	0.5319	0.2785
	總距離	0.5814	0.4052	0.0000	1.9996	1.4392	1.4437	1.7496	1.2887	1.2572

四	下輪廓	0.2542	0.7875	0.5827	0.0000	0.0099	0.0141	0.2125	0.1449	0.1449
	右輪廓	0.4801	0.7182	0.5750	0.0000	0.6108	0.7651	1.0000	0.6042	0.8261
	上輪廓	1.0000	0.7848	0.8420	0.0000	0.0312	0.1852	0.3125	0.3100	0.5635
	總距離	1.7343	2.2905	1.9996	0.0000	0.6519	0.9644	1.5250	1.0592	1.5345
五	下輪廓	0.2642	0.7975	0.5926	0.0099	0.0000	0.0042	0.2025	0.1350	0.1350
	右輪廓	0.1307	0.1074	0.0358	0.6108	0.0000	0.1543	0.3892	0.0066	0.2153
	上輪廓	0.9688	0.7537	0.8108	0.0312	0.0000	0.1540	0.2813	0.2789	0.5323
	總距離	1.3637	1.6585	1.4392	0.6519	0.0000	0.3125	0.8731	0.4205	0.8826
六	下輪廓	0.2684	0.8017	0.5968	0.0141	0.0042	0.0000	0.1983	0.1308	0.1308
	右輪廓	0.2850	0.0469	0.1901	0.7651	0.1543	0.0000	0.2349	0.1609	0.0610
	上輪廓	0.8148	0.5996	0.6568	0.1852	0.1540	0.0000	0.1273	0.1249	0.3783
	總距離	1.3682	1.4482	1.4437	0.9644	0.3125	0.0000	0.5605	0.4166	0.5701
七	下輪廓	0.4667	1.0000	0.7951	0.2125	0.2025	0.1983	0.0000	0.0675	0.0675
	右輪廓	0.5199	0.2818	0.4250	1.0000	0.3892	0.2349	0.0000	0.3958	0.1739
	上輪廓	0.6875	0.4723	0.5295	0.3125	0.2813	0.1273	0.0000	0.0025	0.2510
	總距離	1.6742	1.7541	1.7496	1.5250	0.8731	0.5605	0.0000	0.4658	0.4924
八	下輪廓	0.3992	0.9325	0.7276	0.1449	0.1350	0.1308	0.0675	0.0000	0.0000
	右輪廓	0.1241	0.1140	0.0292	0.6042	0.0066	0.1609	0.3958	0.0000	0.2219
	上輪廓	0.6900	0.4748	0.5319	0.3100	0.2789	0.1249	0.0025	0.0000	0.2535
	總距離	1.2133	1.5213	1.2887	1.0592	0.4205	0.4166	0.4658	0.0000	0.4754
九	下輪廓	0.3992	0.9325	0.7276	0.1449	0.1350	0.1308	0.0675	0.0000	0.0000
	右輪廓	0.3460	0.1079	0.2511	0.8261	0.2153	0.0610	0.1739	0.2219	0.0000
	上輪廓	0.4365	0.2213	0.2785	0.5635	0.5323	0.3783	0.2510	0.2535	0.0000
	總距離	1.1817	1.2617	1.2572	1.5345	0.8826	0.5701	0.4924	0.4754	0.0000

表五、第一種分群

A 群	成員	距離
	一二三	0.9866、0.5814、0.4052
B 群	成員	距離
	四五六	0.6519、0.9644、0.3125
C 群	成員	距離
	七八九	0.4658、0.4924、0.4754

表六、第二種分群

A 群	成員	距離
	一二三	0.9866、0.5814、0.4052
B 群	成員	距離
	四	-
C 群	成員	距離
	五六七八九	0.3125、0.8731、0.4205、0.8826、0.5605、0.4166、0.5701、0.4658、0.4924、0.4754